CSTDM09 - California Statewide Travel Demand Model

Model Development

Model Overview

Final System Documentation: Technical Note

ULTRANS
Institute of Transportation Studies, UC Davis
Davis, California

HBA Specto Incorporated
Calgary, Alberta

May 2011

Table of Contents

Tables:

Figures:

1. Introduction

The 2009 California Statewide Travel Demand Model (CSTDM) forecasts all personal travel made by every California resident, plus all commercial vehicle travel, made on a typical weekday in the fall/spring (when schools are in session). It has five demand models:

- A Short Distance Personal Travel Model (for intra-California trips) (SDPTM);
- A Long Distance Personal Travel Model (for intra-California trips) (LDPTM);
- A Short Distance Commercial Vehicle Model (for intra-California trips) (SDCVM);
- A Long Distance Commercial Vehicle Model (for intra-California trips) (LDCVM);
- An External Vehicle Trip Model (for trips with origin and/or destination outside California).

Figure 1 summarizes the CSTDM09 model system operation.

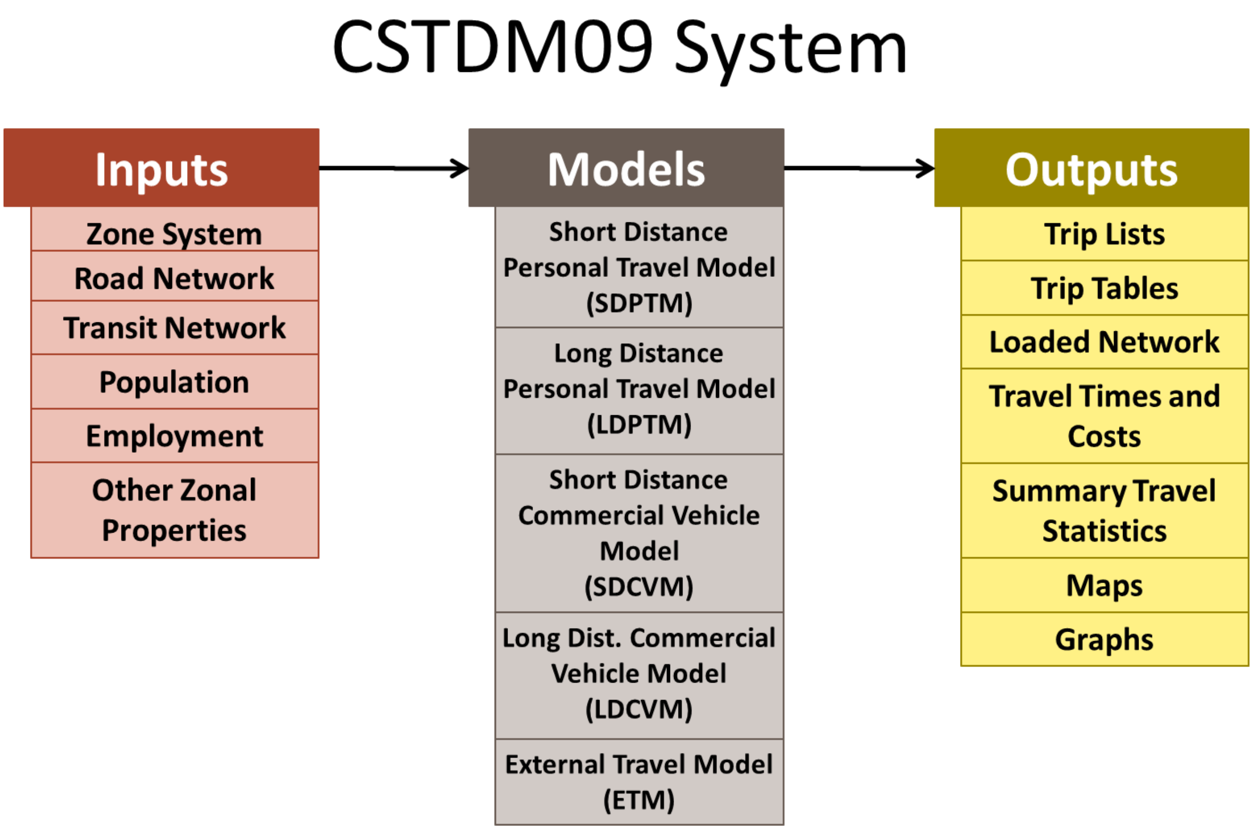

Figure 1: CSTDM09 Model System Operation

The cut-off distance between short and long distance personal travel models is 100 miles, which is defined by the straight-line distance between traffic analysis zone (TAZ) centroids. This 100-mile definition is consistent with that used for the California High Speed Rail Travel Model for person trips. All TAZ-to-TAZ personal travel movements within 100 miles are forecast by the SDPTM; and all TAZ-to-TAZ personal travel movements 100 miles and longer are forecast by the LDPTM.

The cut-off distance between short and long distance commercial vehicle models is 50 miles, which is defined by the straight-line distance between TAZ centroids. This 50-mile definition is consistent with the depot spacing for commercial shippers.

All TAZ-to-TAZ commercial vehicle movements within 50 miles are forecast by the SDCVM; and all TAZ-to-TAZ commercial movements 50 miles and longer are forecast by the LDCVM.

The External Vehicle Trip model forecasts car and commercial vehicle trips made between the 51 external zones and the 5,191 internal TAZs.

The common features of the CSTDM09 models is their disaggregate simulation aspect, their production of a consistent trip list output, and their use of the same inputs where common data is needed (e.g. number of retail employees and travel skims).

2. Inputs

The five demand models use the following inputs:

- Zone system,
- Road network,
- Transit network,
- Population,
- Employment,
- Other zonal properties,

- Estimated model parameters, and
- Calibration factors.

2.1 Zone System

The system of zones used in the CSTDM09 divides California into areas and is consistent with census geography and PECAS land use zones (LUZs). The state is subdivided into 5,191 traffic analysis zones (TAZs). In addition, there are 48 external zone vehicle entry / exit points on roads on the state boundary, plus three external zone seaports whose import / export activities generate significant truck activity (Long Beach, Los Angeles and Oakland). The zones nest both within the 58 California counties and the 524 land use zone (LUZ) system used in the California PECAS spatial economic model. Figure 2 illustrates the CSTDM09 TAZ system and Figure 3 illustrates how TAZs nest within a LUZ with an example in the San Francisco Bay Area.

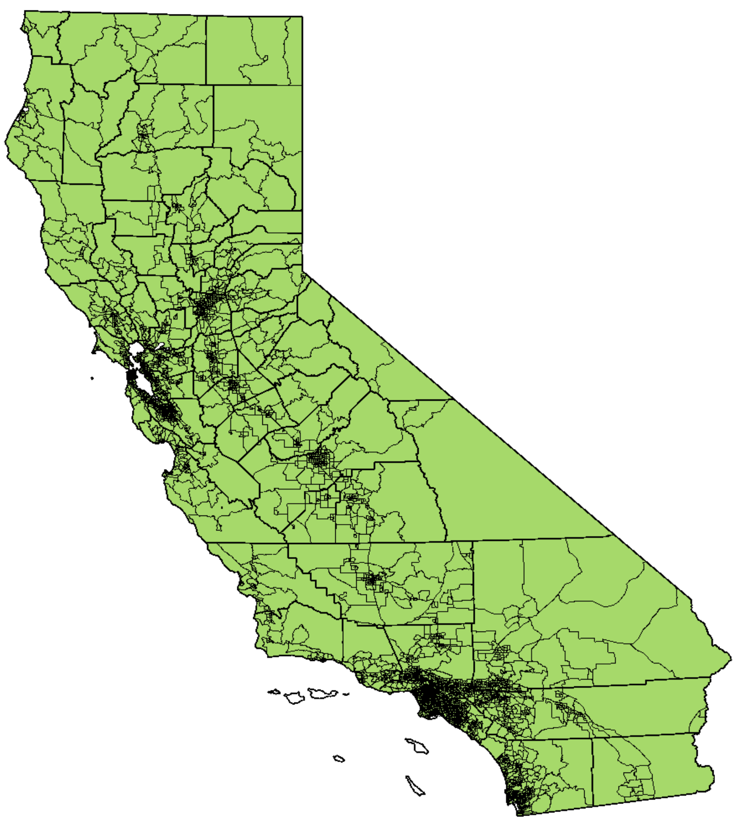

Figure 2: TAZ System

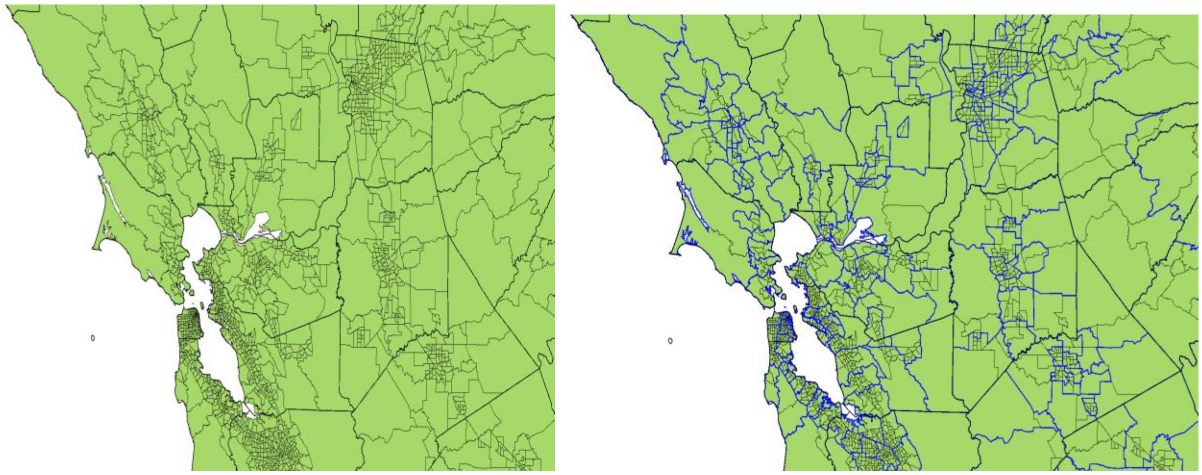

Figure 3: Comparison of TAZ and LUZ in the San Francisco Bay Area

2.2 Networks

2.2.1 Road Networks

The road network represents all freeways, expressways, and most arterial roadways explicitly, with collector and local roads mostly covered through zone centroid connector links. Link distances, free flow speeds and capacities are explicitly coded. Figure 4 illustrates the year 2000 road network coding – the overall network contains over 86,000 nodes and 235,000 links.

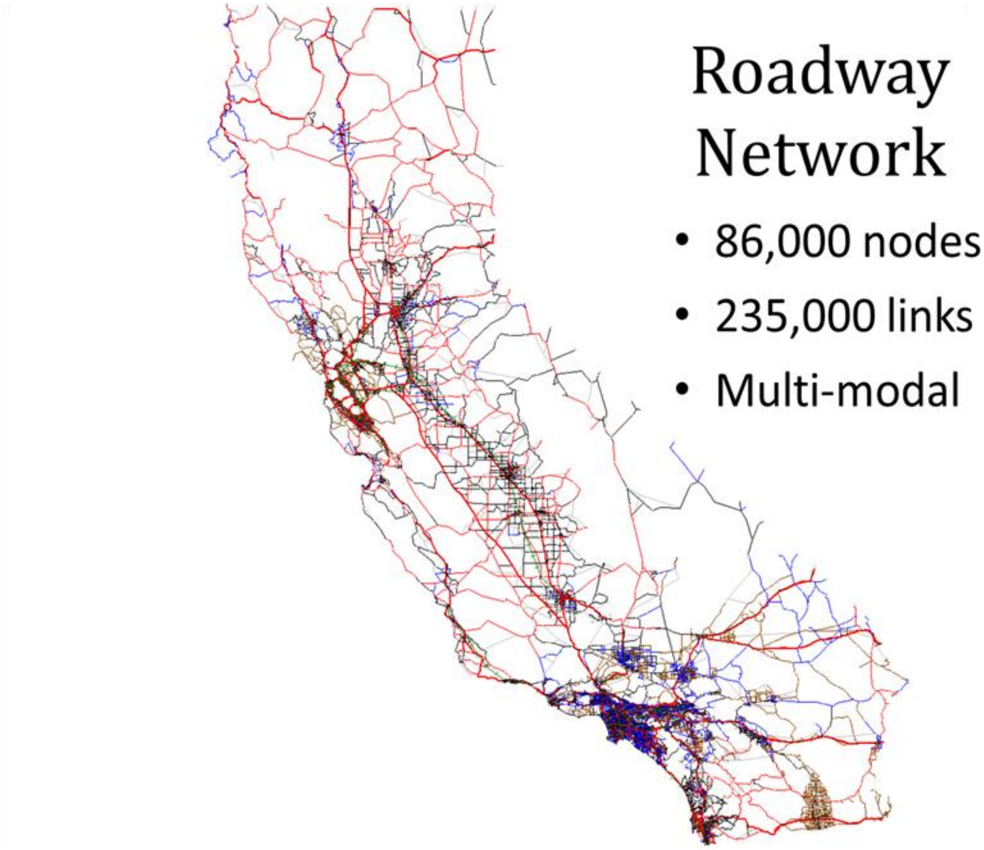

Figure 4: Base Road Network

2.2.2 Transit Network

The transit network combines explicitly coded fixed guideway transit, including all air and rail lines and services, with algorithmically derived local transit (bus). For local bus transit, a simplified model is used to give level of service times and costs, based on road network speeds, land use variables, and transit operator service measures. Observed data (collected through the Google Transit platform) were used to develop the model. The result is reasonable transit representation for a statewide model with a fraction of the coding.

2.2.3 Modes

Each of the travel demand sub-models considers a different set of travel modes relevant to their travel type, as described below and in Table 1.

- The SDPTM considers 8 travel modes – Single occupant vehicle (SOV); high occupant vehicle with 2-persons (HOV2); high occupant vehicle with 3+persons (HOV3); walk access local transit (bus, light rail, heavy rail); drive access local transit (where access to or egress from a rail station is by car); walk; bicycle; and school bus (walk and bicycle times are derived from road network distances).
- The LDPTM considers 5 travel modes – SOV, HOV2, HOV3+, rail, and air.
- The SDCVM considers 3 commercial vehicle types – Light commercial vehicle; medium truck; and heavy truck.
- The LDCVM considers 1 commercial vehicle type – heavy truck.
- The External vehicle Trip model considers 5 travel modes –SOV; HOV2; HOV3; medium truck; and heavy truck.

Table 1: Models in the CSTDM09 Model

Mode	Short Distance Personal	Long Distance Personal	Short Distance Commercial	Long Distance Commercial	External Travel
Auto SOV					
Auto HOV 2 person					
Auto HOV 3+ person					
Transit (bus and rail)					
Bicycle					
Walk					
Air					
Rail					
Light commercial vehicle					
Single unit truck					
Multiple unit truck					

2.2.4 Time Periods

In the CSTMD09, the weekday is split into four time periods for demand modeling and travel assignment purposes, as described in Table 2 below. The models generally further sub-divide the off-peak period into an early time period and a late time period. The early period is defined as being between 3 AM and 6 AM; and the late time period as being between 7 PM and 3 AM. These definitions are consistent with the data collection approach for household travel surveys, where the travel survey day is defined as starting at 3 AM. Road and public transit network descriptions for each time period are coded in the standard CUBE format. Table 2 presents the division of the weekday in the time periods considered in the CSTDM. Please note that Off-peak Early and Off-peak Late are usually combined in the time period “Off Peak” in the CSTDM.

Table 2: Time Periods for the CSTDM09

Start period	Definition
Off-peak Early	3 AM to 6 AM
AM Peak	6 AM to 10AM
Midday	10 AM to 3 PM
PM Peak	3 PM to 7 PM
Off-peak Late	7 PM to 3 AM

2.3 Population

In the CSTDM, the synthetic population represents every person and housing unit in California; it is based on sampling U.S. Census Public Use Microdata Sample (PUMS) 5% person and household data to match targets that can be derived from sources such as Summary File 3 (SF3), the American Community Survey (ACS), or other sources of data. The population synthesizer uses a large number of marginal targets representing categories such as household sizes, housing types, household income groups, person age categories, auto ownership categories, employed workers by occupation category, and students by education level.

2.4 Employment

For the CSTDM09, employment is needed for workers by both industry and occupation. The industry categories describe the type of activity at a person’s place of work, and the occupation categories describe the kind of work a person does to earn a living. For those industries with production and administrative workers, the model accuracy is improved by separating these two types of workers based on occupation information. For instance, a person going to see a movie is interested in going to the nearby cinema, not to Hollywood where the movie is actually produced. For information on industry, we used 13 North American Industrial Classification System (NAICS) categories. NAICS is the federal government's standard industry classification system that groups establishments into industries based on the activities in which they are primarily engaged. 24 Standard Occupational Classification (SOC) categories were used for information on occupation. SOC is the federal government's standard classification system for occupations, and it groups occupations according to the nature of the work performed. An employment synthesizer is used to provide total employment by place of work by finer industry and occupation levels using available datasets from the Census Transportation Planning Package (CTPP), PUMS, California Employment Development Department, and Longitudinal and Household Dynamics (OnTheMap).

2.5 Other Zonal Properties

Other zonal properties include parking costs, area types, districts, and counties.

2.6 Estimated Model Parameters

Estimated model parameters and model forms reflect behavior as estimated in the production of the models from observed behavior (travel surveys).

2.7 Calibration Factors

Calibration factors are adjustments to these parameters to better reflect the observed behavior in the operating model system (travel surveys and count data).

3. Models

3.1 Short Distance Personal Travel Model (SDPTM)

The SDPTM is a disaggregate micro-simulation tour-based choice demand model that was developed for the CSTDM09. Travel Survey data for the years 2000 / 01 from the California State-wide Household Travel Survey and similar surveys conducted in the Southern California Association of Governments, the Metropolitan Transportation Commission, and the San Diego Association of Governments were used to estimate choice-based logit sub-models for components of travel behavior. The model is applied to forecast trips made by every resident of California. Each person / household is assigned to a home TAZ.

The SDPTM is a tour-based travel forecasting model. It uses the concept of a tour as a unit of analysis in the development of model components. A tour represents a closed or half closed chains of trips starting and ending at home or at the workplace. Each tour includes at least one destination and at least two successive trips. A tour is developed by connecting the person trips in a trip chain by time of day, travel activities and stop sequence. Figure 5 illustrates a typical day pattern with three separate tours from / to home; and one sub-tour from / to work.

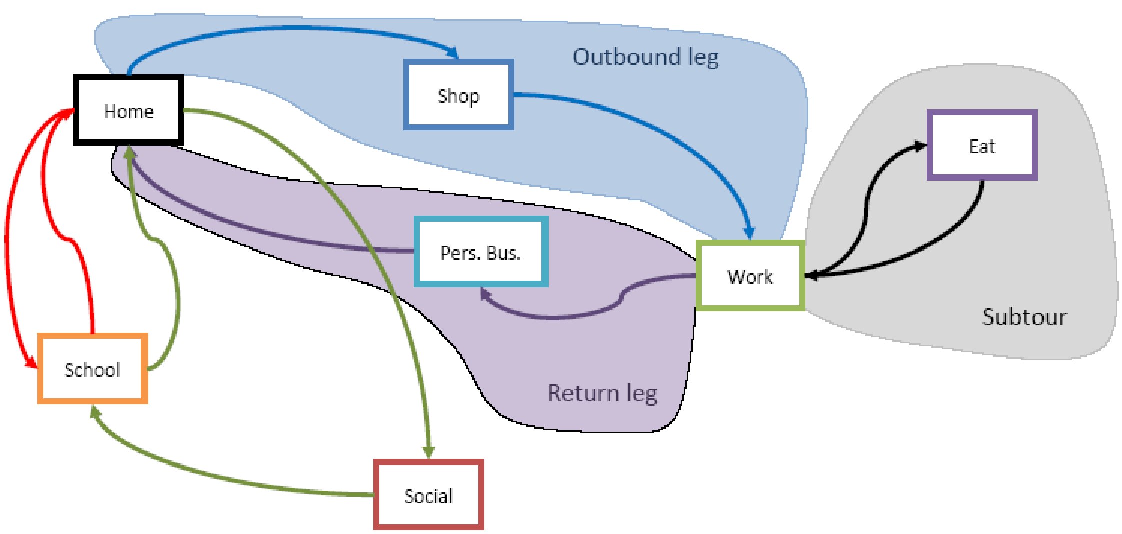

Figure 5: A Typical Day Pattern with Tours

For each tour, a "tour mode" is identified. The tour mode is the overall mode for the tour. The mode that is the "fastest" mode in the trips of a tour is used for the tour mode, defined in a hierarchical order (SOV, HOV2, HOV3+, School Bus, Drive Access Transit, Walk Access Transit, Bicycle, Walk).

The SDPTM has six main components, applied to each person, as shown in Figure 6.

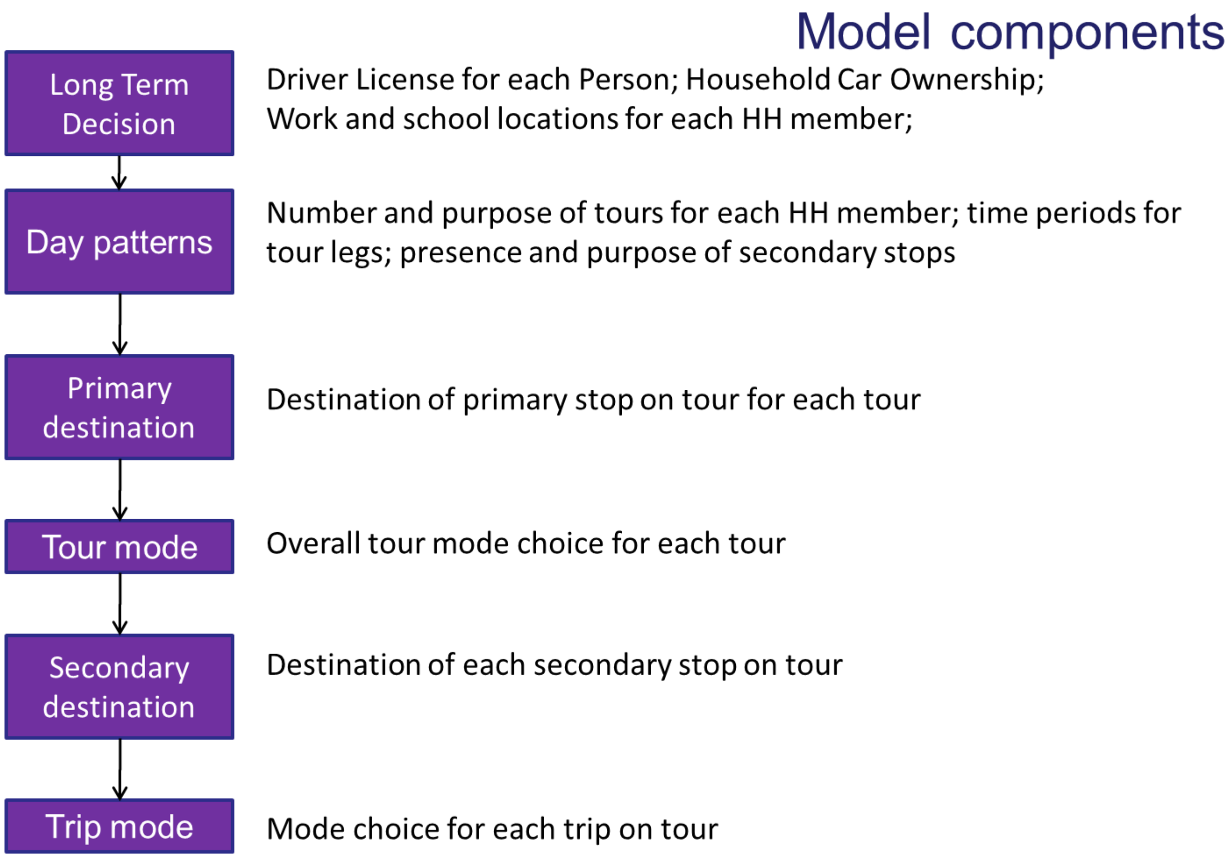

Figure 6: Components of the Short Distance Personal Travel Model

The **Long Term Decision** component of the SDPTM contains the following sub-models for all persons and households:

- A **Driver License** model, which forecasts whether the individual being modeled has a driver license. This model is required because the availability of a driver license is used as an explanatory variable in the household auto ownership and mode choice models of the SDPTM. The “synthetic population” data for each person is obtained from the Federal Census PUMS data, and this data does not contain details of a person’s driving license status.
- A **Household Auto Ownership** model for each individual household, which forecasts 0 to 9 autos for each household. Household Auto availability (defined in 3 categories – 0-auto households, autos < drivers (insufficient), and autos>=

drivers (sufficient)) is an explanatory variable used to forecast mode choice and destination choice (through accessibility measures).

Both the driving license ownership models and the household auto ownership models include demographic and travel “accessibility” explanatory variables. They are thus policy sensitive to change in both demographics and travel “accessibility”.

The **Long Term Decision** component of the SDPTM also contains the following sub-models for all person types:

- A **Work Location** model, which forecasts the potential primary workplace TAZ of the individual being modeled. This location is used as the primary destination for all Work tours made by the individual. (Although this model is used mainly for persons classified as Workers, it is also used to identify primary work locations for other person types, who are forecast to make a work tour as part of their day pattern).
- A **School Location** model, which forecasts the primary school location TAZ of the individual being modeled. This location is used as the primary destination for all School tours made by the individual. (Although this model is primarily used for persons classified as Grade School children or Post-Secondary Education Students, it is also used to identify primary school locations for other person types, who are forecast to make a school tour as part of their day pattern).

The work and school models use a simplified logsum across all possible travel modes to select a work/school location, and they are both sensitive to auto ownership, with the work model also sensitive to income level.

The **Day Pattern** model component of the SDPTM allocates “whole day patterns” for each person, in terms of:

- The number of tours made from home (or the tour start purpose type location if not home);
- The number of stops on each tour, by tour purpose;

- For each tour, the primary tour purpose - defined in a hierarchical fashion:
    - any tour with a Work purpose stop is defined as a Work Tour;
    - any tour with a School purpose stop is defined as a School Tour;
    - the purpose of the first stop for non-Work or School Tours; and
- For each tour, the start and end time periods of the tour.

Note that a tour is generally defined as a set of travel activities to locations other than home, which starts and ends at home. However some tours at the beginning of the travel day can have a start location other than the home – in these cases a tour is defined as complete when the stop location is finally home. Tours not ending at home are treated in a similar fashion.

The **Primary Destination** model component of the SDPTM forecasts the destination of the primary stop on the tour. For Work and School Tours the primary destination has already been forecast by the Long Term Decision Work Location and School Location models. The Primary Destination Models are thus applied for tours where the primary purpose is “Other” i.e. not Work or School. In these cases the primary purpose and destination is defined as the purpose and destination of the first stop.

The **Main Tour Mode** model component of the SDPTM forecasts the main mode used for the tour. This mode is generally used for all trips on the tour, although for certain tour types the **Trip Mode Models** forecast the use of an alternative mode to the main mode.

The **Secondary Destination** model component of the SDPTM forecasts the destination of all secondary stops on the tour, for all tour purposes (Work, School or Other).

The above models are applied differently depending upon whether the tour purpose is Work or School, or whether the tour purpose is Other.

Figure 7 shows the sub-model detail and flow for tours where the tour purpose is Work or School. It has a traditional travel model order, with primary destination chosen before mode choice.

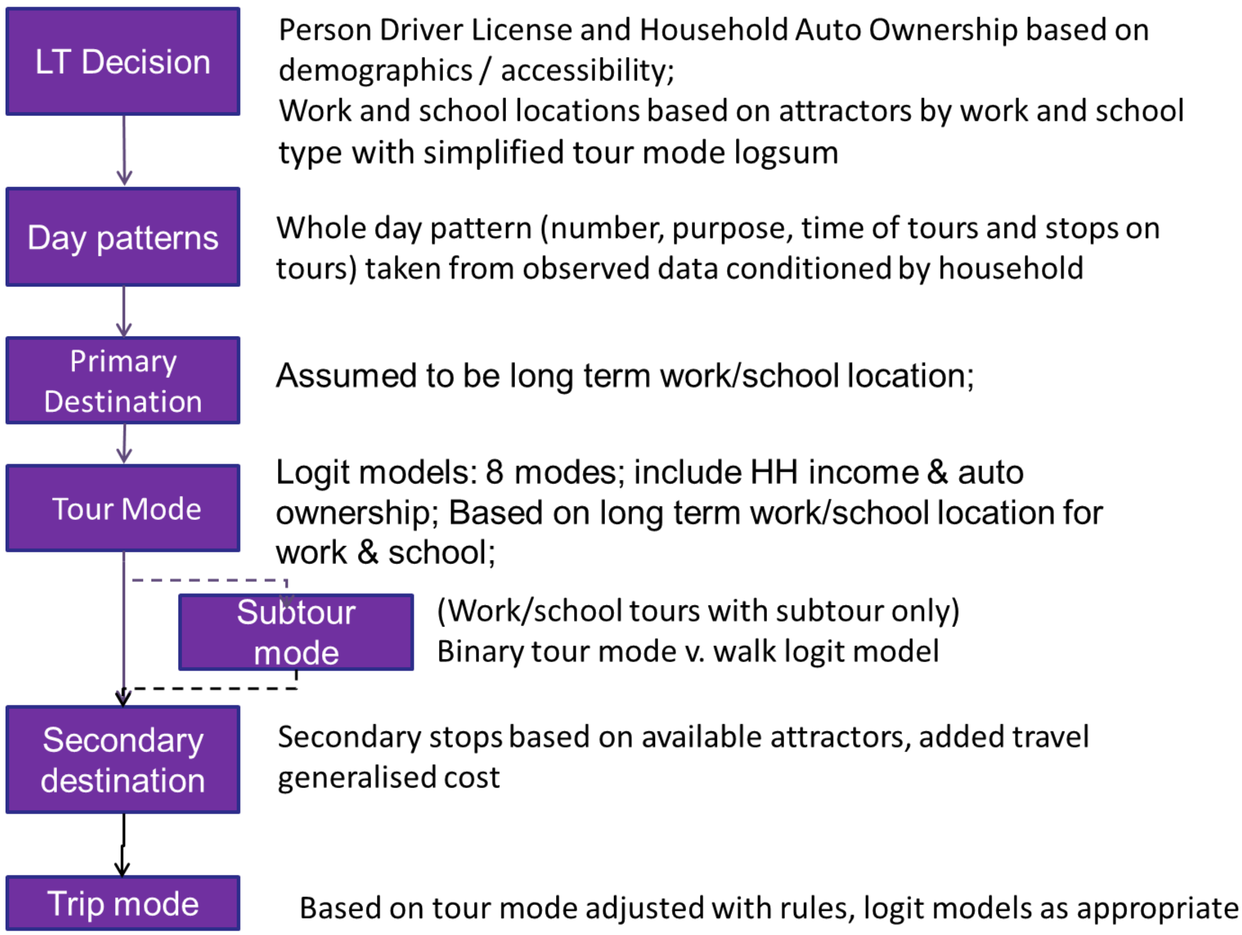

Figure 7: SDPTM Application: Work and School Tours

Figure 8 shows the sub-model detail and flow for tours where the tour purpose is Other. It has a non-traditional travel model order, with tour mode chosen before primary destination.

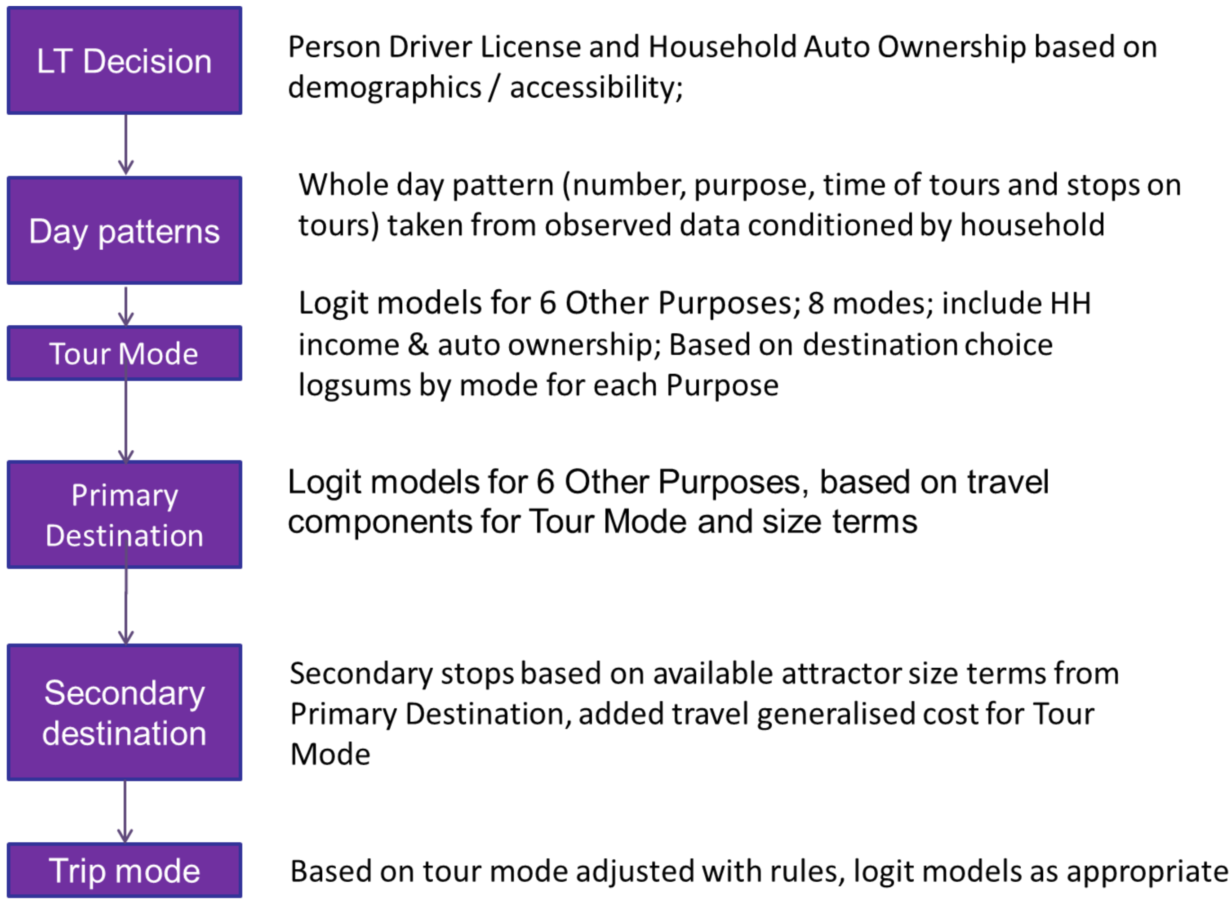

Figure 8: SDPTM Application: Other Tours

3.2 Long Distance Personal Travel Model (LDPTM)

The LDPTM component of the CSTDM framework is based on the Long Distance model component of the HSR model, integrated into the overall CSTDM modeling framework.

The LDPTM is applied at a TAZ level. It predicts trips by each person in the TAZ population, with the following major model components:

1. **Trip Frequency** models for 4 purposes (business; commute; recreation; & other), which predict how many long distance trips for that purpose will be made by each person from & to home;
2. **Party Size** models for 2 purpose groups (business & commute; recreation & other), which predict whether the person is travelling alone or in a group;

3. **Destination Choice** models for 2 purpose groups (business & commute; recreation & other) which predict to or from which TAZ (>100 miles from home TAZ) the long distance trip is made;
4. **Main Mode** models for 2 purpose groups (business & commute; recreation & other), which predict which of 4 potential modes of travel will be chosen for the long distance trip between the home TAZ & the destination zone (Auto; Air; Conventional Rail; High Speed Rail);
5. **Access Mode** models for 2 purpose groups (business & commute; recreation & other), which predict which of 6 potential modes of travel will be chosen for the long distance trip to access the main mode (drive & park; rental car; drop off; taxi; local transit; walk);
6. **Egress Mode** models for 2 purpose groups (business & commute; recreation & other), which predict which of 6 potential modes of travel will be chosen for the long distance trip to egress from the main mode (unpark & drive; rental car; pick up; taxi; local transit; walk); and
7. **Car Occupancy / Time of Day / Direction** factors convert car person trips to car vehicle trips, and split the predicted daily trips by 4 time periods (AM peak, Midday, PM peak and Off-peak) and by direction (from home or to home).

Three sources from the original HSR model work were used in the description of the LDPTM including the:

1. August 2006 HSR Model Development Report,
2. HSR model computer program coded in “Pascal”.
3. “Updated values of Final Coefficients and Constants in the HSR Ridership & Revenue Model” contained in the January 29, 2010 Memorandum from Cambridge Systematics, the HSR Model developer.

3.3 Short Distance Commercial Vehicle Model (SDCVM)

The SDCVM models developed by HBA Specto for the cities of Calgary and Edmonton in Alberta, Canada are being applied in the CSTDM. The models were calibrated using data from Commodity Flow Surveys of over 8,000 business establishments, conducted to determine the characteristics of goods and service movements over a 24-hours period. All sectors of the economy were considered including industrial, wholesale, retail, service, transport and handling and “fleet allocator” (businesses where vehicles operate on regular, and thus relatively fixed, routes rather than making stops in response to individual requirements e.g. parcel delivery / pick-up).

These models are state-of-the-art micro-simulation tour-based models that explicitly predict both goods and service vehicle movements in a local context. They include light, medium, and heavy commercial vehicle movements. They have been implemented in a practical modeling environment. The initial model formulations are based on the Alberta parameter values, and they have been adjusted as appropriate to match California conditions.

The tour-based SDCVM is a group of models that work in series. A basic schematic of the models is shown in Figure 9.

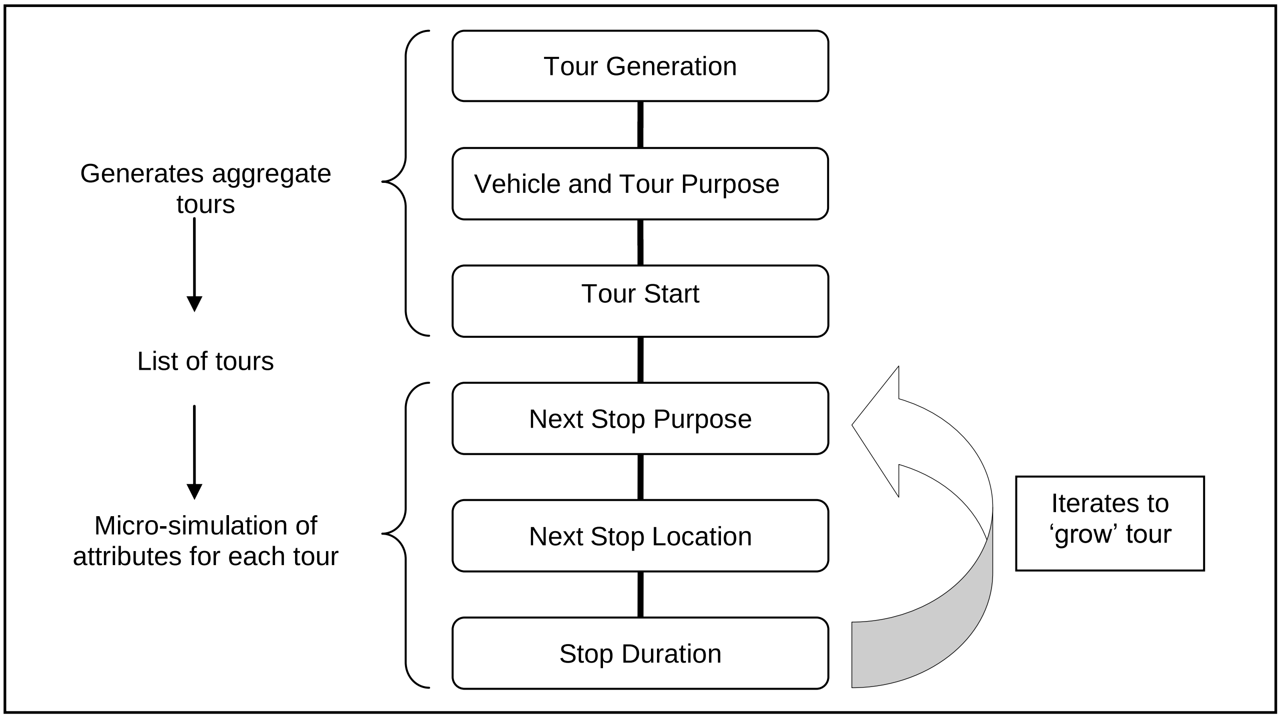

Figure 9: Tour-Based Model Structure

Tour generation quantities by vehicle type, tour purpose, and time of day are generated for each TAZ, using logit and regression equations applied with aggregate TAZ inputs and travel accessibilities, to create a list of tours.

Individual tours generated from each TAZ are then assigned a next stop purpose, next stop location and next stop duration using a micro-simulation process. In this process, Monte Carlo techniques are used to incrementally 'grow' a tour by having a 'return-to-establishment' alternative within the next stop purpose allocation. If the next stop purpose is not 'return-to-establishment', then the tour extends by one more stop. The location and duration of the next stop are then estimated. These steps are repeated until the "return to establishment" next stop purpose is chosen.

Six establishment types are considered, based on aggregations of NAICS categories:

- (Industrial (IN) – NAICS 11, 21, 23, 31-33,;
- Wholesale (WH) – NAICS 42;
- Service (SE) – NAICS 51, 52, 53, 54, 55, 56, 61, 62, 71, 72, 81, 91;
- Retail (RE) – NAICS 44-45;
- Transport and Handling (TH) – NAICS 22, 48-49;
- Fleet Allocator (FA) - All.

Four commercial vehicle types are used:

- Light vehicle FHWA classes 1-3, 5;
- Medium truck < 9.6 short tons – FHWA classes 6-7;
- Medium Truck > 9.6 short tons – FHWA classes 6-7;
- Heavy Truck – FHWA classes 8-13.

3.4 Long Distance Commercial Vehicle Model (LDCVM)

The development of the LDCVM builds directly off the work being done at ULTRANS for the California Department of Transportation (Caltrans), to develop a computer-based model of the California spatial economic system using the PECAS modeling framework. A base year 2000 and 2008 PECAS model is being developed – the same base year used for the state-wide travel model. Output from this PECAS model is used to create an initial year 2000 and 2008 weekday long distance commercial vehicle TAZ to TAZ trip table. Growth factors based on forecast changes in TAZ demographics are then applied to this base commercial vehicle trip table for future year scenarios.

It is important to note that this approach is not dependent upon the availability of future year PECAS model outputs. The derivation of the model uses the base year 2000 and 2008 PECAS model output as input, but application for future year scenarios is carried out using the resulting year commercial vehicle trip table and scaling factors. This means that the travel model can immediately be applied to future year scenarios. In addition, the PECAS model produces truck flows for all zone to zone pairs, for all distance ranges. Only those for origin-destinations greater than 50 miles are applied in the CSTDM.

In the longer term it is expected that future year PECAS model run output will be available directly to inform and enhance the estimation of long distance commercial vehicle flows.

A full description of the PECAS model is given in the documentation of the California PECAS project. A brief overview is given below.

PECAS is a generalized approach for simulating spatial economic systems. It is designed to provide a simulation of the land use component of land use transport interactive modeling systems.

PECAS stands for Production, Exchange, and Consumption Allocation System. Overall, it uses an aggregate, equilibrium structure with separate flows of exchanges (including goods, services, labor and space) going from production to consumption based on variable technical coefficients and market clearing with exchange prices.

It provides an integrated representation of spatially distinct markets for the full range of exchanges, with the transport system and the development of space represented in more detail with specific treatments.

PECAS includes two basic modules that are linked together with travel models and aggregate economic forecasts to provide a representation of the complete spatial economic system.

3.5 External Vehicle Trip Model (ETM)

The ETM is a disaggregate microsimulation model, using exogenous inputs for generation, a logit model for destination choice, and observed shares for the remainder of the aspects of the model. The output of the ETM is a list of trips, in the same format as the other lists of trips produced by the other components of the CSTDM system. Each row in the output file represents a trip, with the various properties (vehicle mode, origin TAZ, destination TAZ, time period, etc). The only difference between the outputs of the ETM and other portions of the CSTDM09 system is that the ETM produces trips where one or both of the origin and destination TAZ are at external stations, where the remainder of the CSTDM produces travel that starts and ends at internal zones.

The ETM has 51 external stations, located at every significant border crossing of California and at the major ports of Los Angeles, Long Beach and Oakland. The 48 road crossings are the same as used in the previous (Dowling Associates) statewide model.

The external stations were classified into six districts; one for crossings on the California/Oregon border, one on the northern part of the California/Nevada border (south to, and including, US Highway 6 near Benton), one for the southern part of the California/Nevada border (starting at State Highway 266 near Oasis), one for the California/Arizona border, one for the California/Mexico border, and one for the ports. These external districts were used for both model preparation and calibration.

4. Model Output

The CSTDM produces consistent travel times and costs for the PECAS model, including: travel benefit measures, network performance measures, and mode-splits by inter and intra regional geographies. The standard outputs of the CSTDM are shown in Figure 10.

During the development of the model, several additional output tables and maps were produced through the screenline tabulation processes (for the comparison of estimated vs. observed traffic volumes on the screenline sections) and post-processor procedures designed for Caltrans.

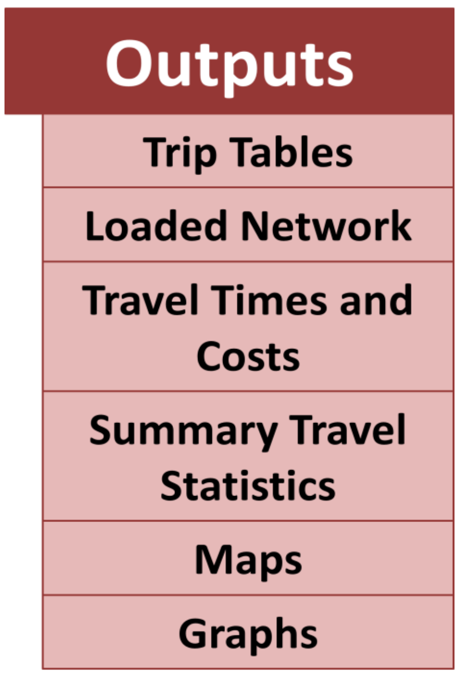

Figure 10: CSTDM09 Model Outputs

In addition, a considerable amount of user-specified output is capable of being produced through CUBE commands, e.g. for volume plots or trip table processing. Depending on the specific needs of the final user, additional output data can be extracted from the demand output tables by specially-written processes using Python, Java, etc.